

0590

10 26

#2 OIPE

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/955,909

DATE: 01/16/2002  
 TIME: 11:05:49

Input Set : N:\Crf3\RULE60\09955909.raw  
 Output Set: N:\CRF3\01162002\I955909.raw

## SEQUENCE LISTING

## 3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: PELLETIER, Marc  
 6 BARKER, William A.  
 7 HAKES, David J.  
 8 ZOPF, David A.

10 (ii) TITLE OF INVENTION: METHODS FOR PRODUCING  
 11 SIALYLOLIGOSACCHARIDES IN A DAIRY SOURCE

13 (iii) NUMBER OF SEQUENCES: 10

15 (iv) CORRESPONDENCE ADDRESS:

16 (A) ADDRESSEE: PENNIE & EDMONDS LLP  
 17 (B) STREET: 1155 Avenue of the Americas  
 18 (C) CITY: New York  
 19 (D) STATE: NY  
 20 (E) COUNTRY: USA  
 21 (F) ZIP: 10036-2711

23 (v) COMPUTER READABLE FORM:

24 (A) MEDIUM TYPE: Diskette  
 25 (B) COMPUTER: IBM Compatible  
 26 (C) OPERATING SYSTEM: DOS  
 27 (D) SOFTWARE: FastSEQ Version 2.0

29 (vi) CURRENT APPLICATION DATA:

C--> 30 (A) APPLICATION NUMBER: US/09/955,909  
 C--> 31 (B) FILING DATE: 18-Sep-2001  
 32 (C) CLASSIFICATION:

34 (vii) PRIOR APPLICATION DATA:

35 (A) APPLICATION NUMBER: 08/911,393  
 36 (B) FILING DATE:

38 (viii) ATTORNEY/AGENT INFORMATION:

39 (A) NAME: Coruzzi, Laura A  
 40 (B) REGISTRATION NUMBER: 30,742  
 41 (C) REFERENCE/DOCKET NUMBER: 7188-032-999

43 (ix) TELECOMMUNICATION INFORMATION:

44 (A) TELEPHONE: (212)7909090  
 45 (B) TELEFAX: (212)8699741  
 46 (C) TELEX: 66141 PENNIE

49 (2) INFORMATION FOR SEQ ID NO: 1:

51 (i) SEQUENCE CHARACTERISTICS:

52 (A) LENGTH: 3183 base pairs  
 53 (B) TYPE: nucleic acid  
 54 (C) STRANDEDNESS: single  
 55 (D) TOPOLOGY: linear

57 (ii) MOLECULE TYPE: cDNA

59 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

61	ATGGGGAAAA	CAGTCGTTGG	GGCCAGTAGG	ATGTTCTGGC	TAATGTTTTT	CGTGCCGCTT	60
62	CTTCTTGCGC	TCTGCCCCAG	CGAGCCCGCG	CATGCCCTGG	CACCCGGATC	GAGCCGAGTT	120
63	GAGCTGTTTA	AGCGGCAAAG	CTCGAAGGTG	CCATTGAAA	AGGGCGGCAA	AGTCACCGAG	180

ENTERED

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/955,909

DATE: 01/16/2002

TIME: 11:05:49

Input Set : N:\Crif3\RULE60\09955909.raw

Output Set: N:\CRF3\01162002\I955909.raw

64	CGGGTTGTCC	ACTCGTTCCG	CCTCCCCGCC	CTTGTTAATG	TGGACGGGGT	GATGGTTGCC	240
65	ATCGCGGACG	CTCGCTACGA	AACATCCAAT	GACAACTCCC	TCATTGATAC	GGTGGCGAAG	300
66	TACAGCGTGG	ACGATGGGGA	GACGTGGGAG	ACCCAAATTG	CCATCAAGAA	CAGTCGTGCA	360
67	TCGTCTGTTT	CTCGTGTGGT	GGATCCCACA	GTGATTGTGA	AGGGCAACAA	GCTTTACGTC	420
68	CTGGTTGGAA	GCTACAACAG	TTCGAGGAGC	TACTGGACGT	CGCATGGTGA	TGCGAGAGAC	480
69	TGGGATATTC	TGCTTGCCGT	TGGTGAGGTC	ACGAAGTCCA	CTGCGGGCGG	CAAGATAACT	540
70	GCGAGTATCA	AATGGGGGAG	CCCCGTGTCA	CTGAAGGAAT	TTTTCCCCGGC	GGAAATGGAA	600
71	GGAAATGCACA	CAAATCAATT	TCTTGCGGGT	GCAGGTGTTG	CCATTGTGGC	GTCCAACGGG	660
72	AATCTTGTGT	ACCCTGTGCA	GGTTACGAAC	AAAAAGAAGC	AAGTTTTTTC	CAAGATCTTC	720
73	TACTCGGAAG	ACGAGGGCAA	GACGTGGAAG	TTTGGGGAGG	GTAGGAGTGA	TTTTGGCTGC	780
74	TCTGAACCTG	TGGCCCTTGA	GTGGGAGGGG	AAGCTCATCA	TAAACACTCG	AGTTGACTAT	840
75	CGCCGCCGTC	TGGTGTACGA	GTCCAGTGAC	ATGGGGAATT	CGTGGGTGGA	GGCTGTCGGC	900
76	ACGCTCTCAC	GTGTGTGGGG	CCCCTCACCA	AAATCGAACC	AGCCCGGCAG	TCAGAGCAGC	960
77	TTCAGTCCG	TGACCATCGA	GGGAATGCGT	GTTATGCTCT	TCACACACCC	GCTGAATTTT	1020
78	AAGGGAAGGT	GGCTGCGCGA	CCGACTGAAC	CTCTGGCTGA	CGGATAACCA	GCGCATTTAT	1080
79	AACGTTGGGC	AAGTATCCAT	TGGTGATGAA	AATTCCGCCT	ACAGTCCCGT	CCTGTACAAG	1140
80	GATGATAAGC	TGTACTGTTT	GCATGAGATC	AACAGTAACG	AGGTGTACAG	CCTTGTTTTT	1200
81	GCGCGCCTGG	TTGGCGAGCT	ACGGATCATT	AAATCAGTGC	TGCAGTCCTG	GAAGAATTGG	1260
82	GACAGCCACC	TGTCCAGCAT	TTGCACCCCT	GCTGATCCAG	CCGCTTCGTC	GTCAGAGCGT	1320
83	GGTTGTGGTC	CCGCTGTAC	CACGGTTGGT	CTTGTTGGCT	TTTTGTGCGA	CAGTGCCACC	1380
84	AAAACCGAAT	GGGAGGATGC	GTACCGCTGC	GTCAACGCAA	GCACGGCAAA	TGCGGAGAGG	1440
85	GTTCCGAACG	GTTTGAAGTT	TGCGGGGGTT	GGCGGAGGGG	CGCTTTGGCC	GGTGAGCCAG	1500
86	CAGGGGCAGA	ATCAACGGTA	TCACTTTGCA	AACCACGCGT	TCACGCTGGT	GGCGTCGGTG	1560
87	ACGATTACAG	AGGTTCCGAG	CGTCGCGAGT	CCTTGCTGG	GTGCGAGCCT	GGACTCTTCT	1620
88	GGTGGCAAAA	AACTCCTGGG	GCTCTCGTAC	GACGAGAAGC	ACCAGTGGCA	GCCAATATAC	1680
89	GGATCAACGC	CGGTGACGCC	GACCGGATCG	TGGGAGATGG	GTAAGAGGTA	CCACGTGGTT	1740
90	CTTACGATGG	CGAATAAAAT	TGGTTCGGTG	TACATTGATG	GAGAACCTCT	GGAGGGTTCA	1800
91	GGGAGACCG	TTGTGCCAGA	CGGGAGGACG	CCTGACATCT	CCCACTTCTA	CGTTGGCGGG	1860
92	TATGGAAGGA	GTGATATGCC	AACCATAAGC	CACGTGACGG	TGAATAATGT	TCTTCTTTAC	1920
93	AACCGTCAGC	TGAATGCCGA	GGAGATCAGG	ACCTTGTTCT	TGAGCCAGGA	CCTGATTGGC	1980
94	ACGGAAGCAC	ACATGGGCAG	CAGCAGCGGC	AGCAGTGCCC	ACAGTACGCC	CTCAACTCCC	2040
95	GCTGACAACG	GTGCCCACAG	TACGCCCTCA	ACTCCGCTG	ACAGCAGTGC	CCACAGTACG	2100
96	CCCTCAACTC	CCGCTGACAG	CAGTGCCAC	AGTACGCCCT	CAGCTCCCGG	TGACAACGGT	2160
97	GCCCACAGTA	CGCCCTCGAC	TCCCGGTGAC	AGCAGTGCCC	ACAGTACGCC	CTCAACTCCC	2220
98	GCTGACAACG	GTGCCCACAG	TACGCCCTCA	GCTCCCGCTG	ACAGCAATGC	CCACAGTACG	2280
99	CCCTCGACTC	CCGCTGACAA	CGGTGCCCAC	AGTACGCCCT	CAACTCCCGC	TGACAACGGT	2340
100	GCCCACAGTA	CGCCCTCGAC	TCCCGGTGAC	AACGGTGCCC	ACAGTACGCC	CTCGACTCCC	2400
101	GGTGACAGCA	GTGCCCACAG	TACGCCCTCA	ACTCCGCTG	ACAACGGTGC	CCACAGTACG	2460
102	CCCTCAGCTC	CCGCTGACAG	CAATGCCCAC	AGTACGCCCT	CGACTCCCGG	TGACAACGGT	2520
103	GCCCACAGTA	CGCCCTCAGC	TCCCGCTGAC	AGCAATGCCC	ACAGTACGCC	CTCGACTCCC	2580
104	GCTGACAGCA	GTGCCCACAG	TACGCCCTCA	GCTCCCGGTG	ACAACGGTGC	CCACAGTACG	2640
105	CCCTCAGCTC	CCGCTGACAG	CAGTGCCCAC	AGTACGCCCT	CAGCTCCCGG	TGACAACGGT	2700
106	GCCCACAGTA	CGCCCTCAGC	TCCCGCTGAC	AACGGTGCCC	ACAGTACGCC	CTCAGCTCCC	2760
107	GGTGACAGCA	ATGCCCACAG	TACGCCCTCG	ACTCCGCTG	ACAGCAGTGC	CCACAGTACG	2820
108	CCCTCAACTC	CCGCTGACAG	CAGTGCCCAC	AGTACGCCCT	CAGCTCCCGG	TGACAACGGT	2880
109	GCCCACAGTA	CGCCCTCAGC	TCCCGCTGAC	AGCAGTGCCC	ACAGTACGCC	CTCAATTCCC	2940
110	GGTGACAGCA	GTGCCCACAG	TACGCCCTCA	GCTCCCGCTG	ACAGCAGTGC	CCACAGTACG	3000
111	CCCTCAGCTC	CCGCTGACAA	CGGTGCCCAC	AGTACGCCCT	CGACTCCCGC	TGACAACGGC	3060
112	GCTAATGGTA	CGGTTTTGAT	TTTGACGAT	GGCGCTGCAT	TTTCGGCCTT	TTCGGGCGGA	3120

## RAW SEQUENCE LISTING

DATE: 01/16/2002

PATENT APPLICATION: US/09/955,909

TIME: 11:05:49

Input Set : N:\Crif3\RULE60\09955909.raw

Output Set: N:\CRF3\01162002\I955909.raw

```

113 GGGCTTCTTT TGTGTGCGGG TGCTTTGCTG CTGCACGTGT TCGTTATGGC AGTTTTTTTC 3180
114 TGA 3183
116 (2) INFORMATION FOR SEQ ID NO: 2:
118 (i) SEQUENCE CHARACTERISTICS:
119 (A) LENGTH: 1060 amino acids
120 (B) TYPE: amino acid
121 (C) STRANDEDNESS: single
122 (D) TOPOLOGY: linear
124 (ii) MOLECULE TYPE: protein
126 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
128 Met Gly Lys Thr Val Val Gly Ala Ser Arg Met Phe Trp Leu Met Phe
129 1 5 10 15
130 Phe Val Pro Leu Leu Leu Ala Leu Cys Pro Ser Glu Pro Ala His Ala
131 20 25 30
132 Leu Ala Pro Gly Ser Ser Arg Val Glu Leu Phe Lys Arg Gln Ser Ser
133 35 40 45
134 Lys Val Pro Phe Glu Lys Gly Gly Lys Val Thr Glu Arg Val Val His
135 50 55 60
136 Ser Phe Arg Leu Pro Ala Leu Val Asn Val Asp Gly Val Met Val Ala
137 65 70 75 80
138 Ile Ala Asp Ala Arg Tyr Glu Thr Ser Asn Asp Asn Ser Leu Ile Asp
139 85 90 95
140 Thr Val Ala Lys Tyr Ser Val Asp Asp Gly Glu Thr Trp Glu Thr Gln
141 100 105 110
142 Ile Ala Ile Lys Asn Ser Arg Ala Ser Ser Val Ser Arg Val Val Asp
143 115 120 125
144 Pro Thr Val Ile Val Lys Gly Asn Lys Leu Tyr Val Leu Val Gly Ser
145 130 135 140
146 Tyr Asn Ser Ser Arg Ser Tyr Trp Thr Ser His Gly Asp Ala Arg Asp
147 145 150 155 160
148 Trp Asp Ile Leu Leu Ala Val Gly Glu Val Thr Lys Ser Thr Ala Gly
149 165 170 175
150 Gly Lys Ile Thr Ala Ser Ile Lys Trp Gly Ser Pro Val Ser Leu Lys
151 180 185 190
152 Glu Phe Phe Pro Ala Glu Met Glu Gly Met His Thr Asn Gln Phe Leu
153 195 200 205
154 Gly Gly Ala Gly Val Ala Ile Val Ala Ser Asn Gly Asn Leu Val Tyr
155 210 215 220
156 Pro Val Gln Val Thr Asn Lys Lys Lys Gln Val Phe Ser Lys Ile Phe
157 225 230 235 240
158 Tyr Ser Glu Asp Glu Gly Lys Thr Trp Lys Phe Gly Glu Gly Arg Ser
159 245 250 255
160 Asp Phe Gly Cys Ser Glu Pro Val Ala Leu Glu Trp Glu Gly Lys Leu
161 260 265 270
162 Ile Ile Asn Thr Arg Val Asp Tyr Arg Arg Arg Leu Val Tyr Glu Ser
163 275 280 285
164 Ser Asp Met Gly Asn Ser Trp Val Glu Ala Val Gly Thr Leu Ser Arg
165 290 295 300
166 Val Trp Gly Pro Ser Pro Lys Ser Asn Gln Pro Gly Ser Gln Ser Ser

```

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/955,909

DATE: 01/16/2002  
 TIME: 11:05:49

Input Set : N:\Crf3\RULE60\09955909.raw  
 Output Set: N:\CRF3\01162002\I955909.raw

167	305				310				315				320			
168	Phe	Thr	Ala	Val	Thr	Ile	Glu	Gly	Met	Arg	Val	Met	Leu	Phe	Thr	His
169					325					330					335	
170	Pro	Leu	Asn	Phe	Lys	Gly	Arg	Trp	Leu	Arg	Asp	Arg	Leu	Asn	Leu	Trp
171				340					345					350		
172	Leu	Thr	Asp	Asn	Gln	Arg	Ile	Tyr	Asn	Val	Gly	Gln	Val	Ser	Ile	Gly
173			355					360					365			
174	Asp	Glu	Asn	Ser	Ala	Tyr	Ser	Ser	Val	Leu	Tyr	Lys	Asp	Asp	Lys	Leu
175		370					375					380				
176	Tyr	Cys	Leu	His	Glu	Ile	Asn	Ser	Asn	Glu	Val	Tyr	Ser	Leu	Val	Phe
177	385					390					395				400	
178	Ala	Arg	Leu	Val	Gly	Glu	Leu	Arg	Ile	Ile	Lys	Ser	Val	Leu	Gln	Ser
179				405					410						415	
180	Trp	Lys	Asn	Trp	Asp	Ser	His	Leu	Ser	Ser	Ile	Cys	Thr	Pro	Ala	Asp
181			420						425					430		
182	Pro	Ala	Ala	Ser	Ser	Ser	Glu	Arg	Gly	Cys	Gly	Pro	Ala	Val	Thr	Thr
183			435					440					445			
184	Val	Gly	Leu	Val	Gly	Phe	Leu	Ser	His	Ser	Ala	Thr	Lys	Thr	Glu	Trp
185		450				455					460					
186	Glu	Asp	Ala	Tyr	Arg	Cys	Val	Asn	Ala	Ser	Thr	Ala	Asn	Ala	Glu	Arg
187	465					470					475				480	
188	Val	Pro	Asn	Gly	Leu	Lys	Phe	Ala	Gly	Val	Gly	Gly	Gly	Ala	Leu	Trp
189				485					490					495		
190	Pro	Val	Ser	Gln	Gly	Gln	Asn	Gln	Arg	Tyr	His	Phe	Ala	Asn	His	
191			500					505					510			
192	Ala	Phe	Thr	Leu	Val	Ala	Ser	Val	Thr	Ile	His	Glu	Val	Pro	Ser	Val
193			515					520					525			
194	Ala	Ser	Pro	Leu	Leu	Gly	Ala	Ser	Leu	Asp	Ser	Ser	Gly	Gly	Lys	Lys
195		530					535					540				
196	Leu	Leu	Gly	Leu	Ser	Tyr	Asp	Glu	Lys	His	Gln	Trp	Gln	Pro	Ile	Tyr
197	545					550					555				560	
198	Gly	Ser	Thr	Pro	Val	Thr	Pro	Thr	Gly	Ser	Trp	Glu	Met	Gly	Lys	Arg
199				565					570					575		
200	Tyr	His	Val	Val	Leu	Thr	Met	Ala	Asn	Lys	Ile	Gly	Ser	Val	Tyr	Ile
201				580					585					590		
202	Asp	Gly	Glu	Pro	Leu	Glu	Gly	Ser	Gly	Gln	Thr	Val	Val	Pro	Asp	Gly
203			595					600					605			
204	Arg	Thr	Pro	Asp	Ile	Ser	His	Phe	Tyr	Val	Gly	Gly	Tyr	Gly	Arg	Ser
205		610					615					620				
206	Asp	Met	Pro	Thr	Ile	Ser	His	Val	Thr	Val	Asn	Asn	Val	Leu	Leu	Tyr
207	625					630					635				640	
208	Asn	Arg	Gln	Leu	Asn	Ala	Glu	Glu	Ile	Arg	Thr	Leu	Phe	Leu	Ser	Gln
209				645					650					655		
210	Asp	Leu	Ile	Gly	Thr	Glu	Ala	His	Met	Gly	Ser	Ser	Ser	Gly	Ser	Ser
211				660					665					670		
212	Ala	His	Ser	Thr	Pro	Ser	Thr	Pro	Ala	Asp	Asn	Gly	Ala	His	Ser	Thr
213			675					680					685			
214	Pro	Ser	Thr	Pro	Ala	Asp	Ser	Ser	Ala	His	Ser	Thr	Pro	Ser	Thr	Pro
215		690						695				700				

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/955,909

DATE: 01/16/2002  
 TIME: 11:05:49

Input Set : N:\Crf3\RULE60\09955909.raw  
 Output Set: N:\CRF3\01162002\I955909.raw

```

216 Ala Asp Ser Ser Ala His Ser Thr Pro Ser Ala Pro Gly Asp Asn Gly
217 705 710 715 720
218 Ala His Ser Thr Pro Ser Thr Pro Gly Asp Ser Ser Ala His Ser Thr
219 725 730 735
220 Pro Ser Thr Pro Ala Asp Asn Gly Ala His Ser Thr Pro Ser Ala Pro
221 740 745 750
222 Ala Asp Ser Asn Ala His Ser Thr Pro Ser Thr Pro Ala Asp Asn Gly
223 755 760 765
224 Ala His Ser Thr Pro Ser Thr Pro Ala Asp Asn Gly Ala His Ser Thr
225 770 775 780
226 Pro Ser Thr Pro Gly Asp Asn Gly Ala His Ser Thr Pro Ser Thr Pro
227 785 790 795 800
228 Gly Asp Ser Ser Ala His Ser Thr Pro Ser Thr Pro Ala Asp Asn Gly
229 805 810 815
230 Ala His Ser Thr Pro Ser Ala Pro Ala Asp Ser Asn Ala His Ser Thr
231 820 825 830
232 Pro Ser Thr Pro Gly Asp Asn Gly Ala His Ser Thr Pro Ser Ala Pro
233 835 840 845
234 Ala Asp Ser Asn Ala His Ser Thr Pro Ser Thr Pro Ala Asp Ser Ser
235 850 855 860
236 Ala His Ser Thr Pro Ser Ala Pro Gly Asp Asn Gly Ala His Ser Thr
237 865 870 875 880
238 Pro Ser Ala Pro Ala Asp Ser Ser Ala His Ser Thr Pro Ser Ala Pro
239 885 890 895
240 Gly Asp Asn Gly Ala His Ser Thr Pro Ser Ala Pro Ala Asp Asn Gly
241 900 905 910
242 Ala His Ser Thr Pro Ser Ala Pro Gly Asp Ser Asn Ala His Ser Thr
243 915 920 925
244 Pro Ser Thr Pro Ala Asp Ser Ser Ala His Ser Thr Pro Ser Thr Pro
245 930 935 940
246 Ala Asp Ser Ser Ala His Ser Thr Pro Ser Ala Pro Gly Asp Asn Gly
247 945 950 955 960
248 Ala His Ser Thr Pro Ser Ala Pro Ala Asp Ser Ser Ala His Ser Thr
249 965 970 975
250 Pro Ser Ile Pro Gly Asp Ser Ser Ala His Ser Thr Pro Ser Ala Pro
251 980 985 990
252 Ala Asp Ser Ser Ala His Ser Thr Pro Ser Ala Pro Gly Asp Asn Gly
253 995 1000 1005
254 Ala His Ser Thr Pro Ser Thr Pro Ala Asp Asn Gly Ala Asn Gly Thr
255 1010 1015 1020
256 Val Leu Ile Leu His Asp Gly Ala Ala Phe Ser Ala Phe Ser Gly Gly
257 1025 1030 1035 1040
258 Gly Leu Leu Leu Cys Ala Gly Ala Leu Leu Leu His Val Phe Val Met
259 1045 1050 1055
260 Ala Val Phe Phe
261 1060
263 (2) INFORMATION FOR SEQ ID NO: 3:
265 (i) SEQUENCE CHARACTERISTICS:
266 (A) LENGTH: 1929 base pairs

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/955,909

DATE: 01/16/2002

TIME: 11:05:50

Input Set : N:\Crf3\RULE60\09955909.raw

Output Set: N:\CRF3\01162002\I955909.raw

L:30 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]  
L:31 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]